Introduction to Probability and Statistics

Sec 3: Continuous Random Variables

Random Variables (Recap)

A **Random Variable (RV)** is a variable which assumes numerical values associated with the random outcomes of an experiment.

One, and only one, numerical value is assigned to each outcome of the experiment.

In the previous unit, we studied discrete random variables. We will now turn our attention to Continuous Random Variables.

For example:

- Highest daily temperature in Dublin.
- Exchange rate of EUR to USD.
- Time between two consecutive calls received at a call centre.

How did we Proceed for Discrete RVs?

1. **Define general concepts and formulas**
   - Probability Mass Function
   - Cumulative Distribution Function
   - Expectation and its computation (weighted average)
   - Variance and its computation (also weighted average)

2. **Studied a few classical examples of discrete distributions**
   - Bernoulli
   - Binomial
   - Poisson

We'll do the same for Continuous RVs

Discrete		Continuous
Probability mass function	$\longrightarrow$	Probability density function
Cumulative distribution function	$\longrightarrow$	Cumulative distribution function
Expectation and Variance	$\longrightarrow$	Expectation and Variance
Finite sums: $\sum_i$	*In general* $\longrightarrow$	Integrals: $\int \ldots dx$

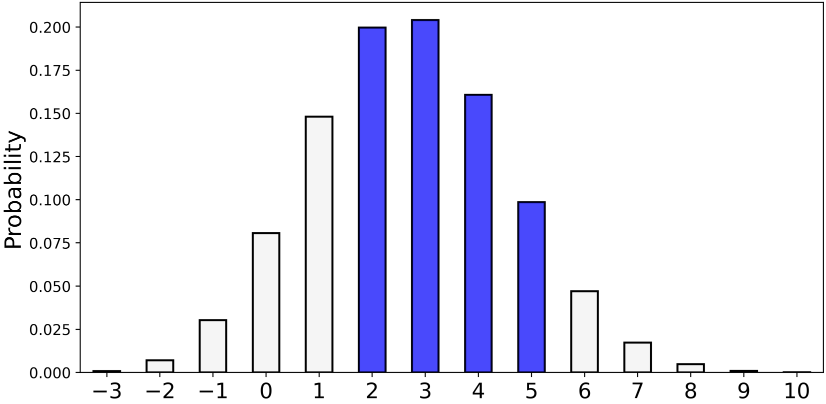
We will then study in more detail the properties of some classical continuous distributions, such as Uniform, Exponential, and Normal.

Continuous Random Variables

- A continuous random variable X can assume an *infinite* number of values, usually all values in one interval $\mathcal{I}$ (could be $\mathcal{I} = \mathbb{R}, \mathbb{R}^+, [a, b], \ldots$).
- As such, the prob. that X is exactly equal to any one specific number is zero!
- Think of $X =$ time it takes you from the moment you enter until you sit.
  - What's the probability it takes you *exactly* 14.3845251 seconds? 0!
  - What's the probability it takes you *exactly* 15 seconds? 0!
- Instead, it is meaningful to ask:
  - What's the probability it takes you between 10 and 15 seconds?
- So, for a continuous rv, we want to be able to compute the probability that X falls in a given *interval*: $\mathbb{P}(x_0 \leq X \leq x_1)$. Not that $X =$ one specific number.

One Step Back

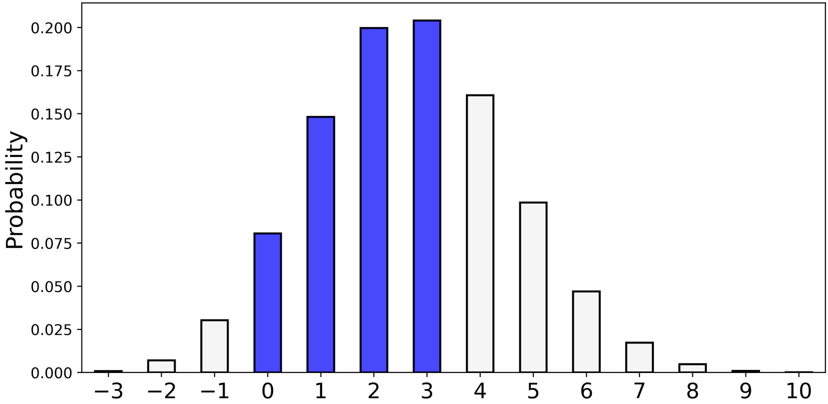
Even for discrete distributions, we could (and did) consider probability of intervals



$$P(2 \leq X \leq 5)$$

One Step Back

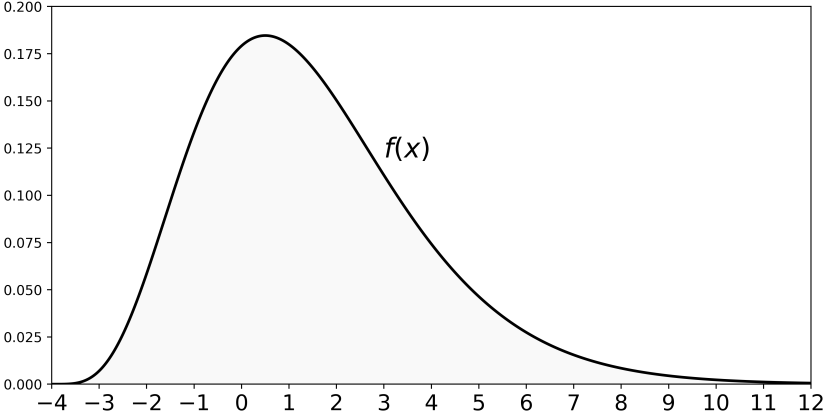
Even for discrete distributions, we could (and did) consider probability of intervals



$$P(0 \leq X \leq 3)$$

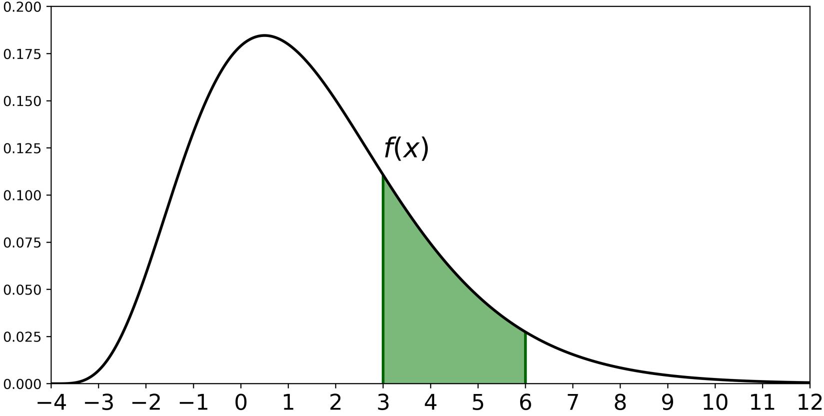
Equivalent for Continuous RVs: Density Function

For continuous random variables, we have a probability density function $f(x)$



Equivalent for Continuous RVs: Density Function

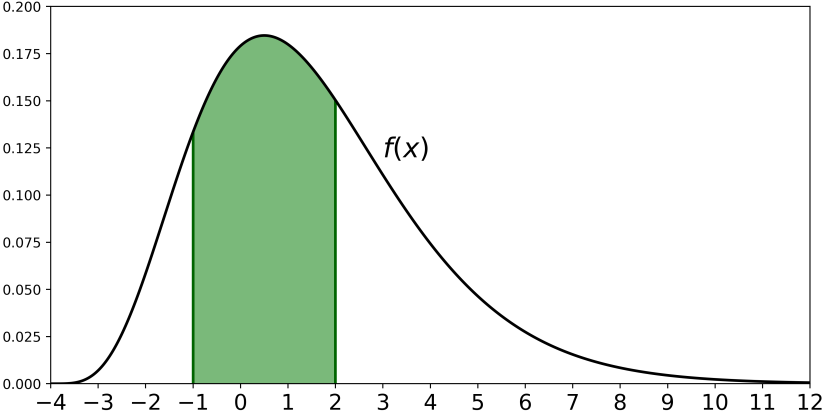
For continuous random variables, we have a probability density function $f(x)$



$$P(3 \leq X \leq 6)$$

Equivalent for Continuous RVs: Density Function

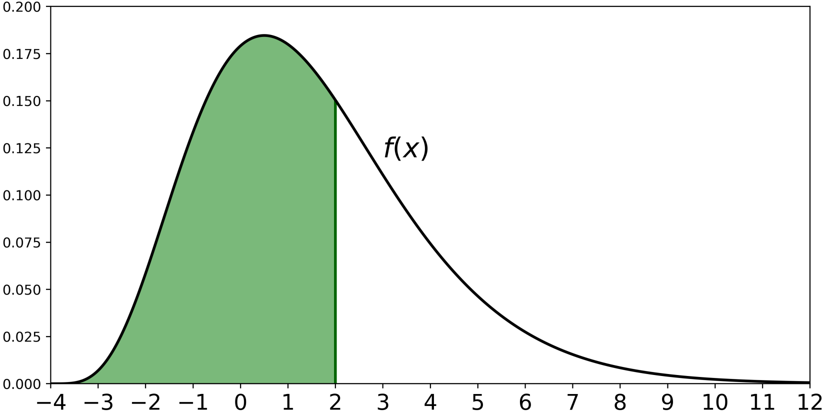
For continuous random variables, we have a probability density function $f(x)$



$P(-1 \leq X \leq 2)$

Equivalent for Continuous RVs: Density Function

For continuous random variables, we have a probability density function $f(x)$



$P(X \leq 2)$

Probability Density Function

- We specify the distribution of a continuous random variable X via its **probability density function (pdf)** $f(x)$.

- For a function $f(x)$ to be a pdf, we need two conditions:

  1. $f(x) \geq 0 \quad \forall x$
  2. $\displaystyle\int_{-\infty}^{+\infty} f(x)\, dx = 1$

- The concept of **pdf** $f(x)$ for continuous RVs is equivalent to the one of **pmf** $p(x)$ for discrete RVs. Recall also how, in the discrete case, we had:

  1. $p(x) \geq 0 \quad \forall x$
  2. $\displaystyle\sum_x p(x) = 1$

Probability Density Function

- If X has density $f(x)$, then the probability that X takes a value between two points x_0 and x_1 is given by the area under the density curve between these two points. This is the integral between the two points.

$$\mathbb{P}(x_0 < X < x_1) = \int_{x_0}^{x_1} f(x)dx$$

- Since there is no area over a single point:

$$\mathbb{P}(X = x) = 0 \quad \forall x \in \mathbb{R}.$$

- Since single points do not contribute to the probability, we also have:

$$\mathbb{P}(x_0 < X < x_1) = \mathbb{P}(x_0 \leq X \leq x_1).$$

Cumulative Distribution Function

- As in the discrete case, we can consider cumulative probabilities:

$$F(x) = \mathbb{P}(X \leq x).$$

  Once again, $F(x)$ is called the cumulative distribution function of X.

- Graphically, $F(x)$ is the area under the density to the left of a point x.

- It is found by integrating the density function between the *lower limit of the values X can take*, and x.

$$F(x) \;=\; \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(u)du$$

  Note: there may be a more meaningful choice than $-\infty$ for the lower limit. Eg, 0 if $X \geq 0$.

Expected Value

- To find the **expected value** of X, we need to compute the "weighted average" of all values X can take, weighing them by how "likely" they are.

- Mathematically, this is an integral, where the weights are given by $f(x)$.

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

Note: you can limit the bounds of integration only to the range of X.

- Once again, notice the analogy with the discrete case: sum replaced by integral, pmf $p(x)$ replaced by pdf $f(x)$.

$$\mathbb{E}[X] = \sum_{x} x\, p(x)$$

Expected Value of a function of X

- More generally, if we want to compute the expected value of some function of X, $g(X)$, then we still weigh the values of $g(X)$ by the density of X. So:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x)\, f(x)\, dx$$

- For example, the expected value of X^2 will be:

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{+\infty} x^2\, f(x)\, dx$$

Variance

- Since the variance is still an expected value (the one of $(X - \mathbb{E}[X])^2$), the formula for the variance follows from before:

$$\mathbb{V}\text{ar}[X] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)\, dx, \quad \text{where } \mu = \mathbb{E}[X].$$

- As in the discrete case, an alternative and sometimes more convenient formula to compute the variance is:

$$\mathbb{V}\text{ar}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2.$$

Well-known Continuous Distributions

In the rest of the chapter, we will study the following examples of continuous distributions:

- Uniform
- Exponential
- Normal (or Gaussian)

Note that there are many more examples of "common" continuous distributions, *e.g.*, Beta, Gamma, Weibull, Cauchy, Chi squared...

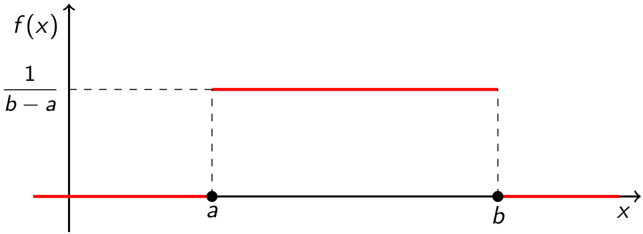
Uniform Distribution

The Uniform Distribution

- A **uniform random variable** over an interval $[a, b]$ is a continuous RV for which all outcomes in the range $[a, b]$ are equally likely. We write:

$$X \sim \mathcal{U}(a, b).$$

- By definition, the **density** of a $X \sim \mathcal{U}(a, b)$ is "flat" over $[a, b]$, and 0 outside the interval:



The Uniform Distribution

- A **uniform random variable** over an interval $[a, b]$ is a continuous RV for which all outcomes in the range $[a, b]$ are equally likely. We write:

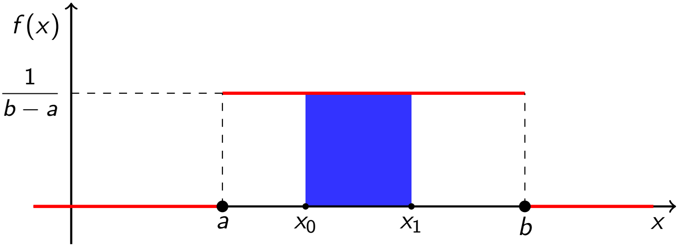
$$X \sim \mathcal{U}(a, b).$$

- By definition, the **density** of a $X \sim \mathcal{U}(a, b)$ is "flat" over $[a, b]$, and 0 outside the interval:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Probability of an Interval

For any interval $[x_0, x_1]$ (contained in $[a, b]$), the probability of the interval is found by computing the area of the associated rectangle:



- This has width $x_1 - x_0$ and height $1/(b-a)$. So:

$$\mathbb{P}(x_0 \leq X \leq x_1) = \frac{x_1 - x_0}{b - a}.$$

CDF and Other Properties

- The **cumulative distribution function** for $X \sim \mathcal{U}(a, b)$ is:

$$F(x) = \mathbb{P}(X \leq x) = \begin{cases} 0 & \text{if } x < a \\ \dfrac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$$

This follows from the previous slide, by considering the interval $[a, x]$.

- **Mean and Variance**:

$$\mathbb{E}[X] = \frac{a+b}{2}, \qquad \mathbb{V}\mathrm{ar}\,[X] = \frac{(b-a)^2}{12}.$$

Uniform Distribution: Example

A company's rolling machine is producing sheets of steel of varying thickness. The thickness T is known to take values uniformly between 1.5 and 2 mm. Sheets less than 1.6 mm thick must be scrapped.

1. What proportion of sheets do you expect to be scrapped?
2. Calculate the mean and standard deviation of the sheet thickness T.
3. Graph the probability density function of T. On the x axis, mark the mean as well as the 2-std wide interval centered in the mean (1 std to the right and 1 std to the left of the mean).

Uniform Distribution: Example

A company's rolling machine is producing sheets of steel of varying thickness. The thickness T is known to take values uniformly between 1.5 and 2 mm. Sheets less than 1.6 mm thick must be scrapped.

1. What proportion of sheets do you expect to be scrapped?

$$\mathbb{P}(T \leq 1.6) = \frac{1.6 - 1.5}{2 - 1.5} = \frac{0.1}{0.5} = 0.2.$$

2. Calculate the mean and standard deviation of the sheet thickness T.

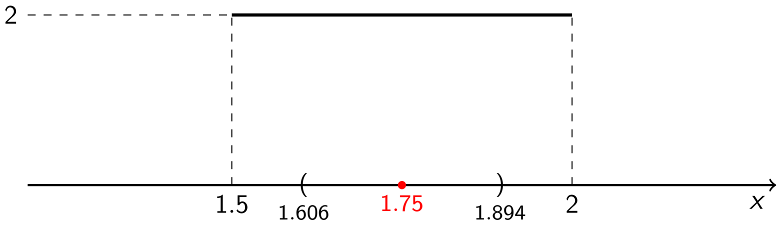
$$\mathbb{E}[T] = \frac{1.5 + 2}{2} = 1.75$$

$$\mathbb{V}\mathrm{ar}[T] = \frac{(2 - 1.5)^2}{12} = 0.02083 \quad \Rightarrow \quad \mathrm{Std}[T] = 0.144\,.$$

Uniform Distribution: Example

A company's rolling machine is producing sheets of steel of varying thickness. The thickness T is known to take values uniformly between 1.5 and 2 mm. Sheets less than 1.6 mm thick must be scrapped.

3. Graph the probability density function of T. On the x axis, mark the mean as well as the 2-std wide interval centered in the mean (1 std to the right and 1 std to the left of the mean).



Exponential Distribution

Exponential Distribution

The **exponential distribution** is typically used to model *waiting times*. E.g.:

- Time between server failures.
- Time spent waiting for a train.

We'll see why that is the case. For now, let's start with seeing what the shape of the exponential density looks like and the properties of the exponential distribution.

Density

The density of an Exponential random variable is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases},$$

where the parameter $\lambda > 0$. We write: $X \sim Exp(\lambda)$.

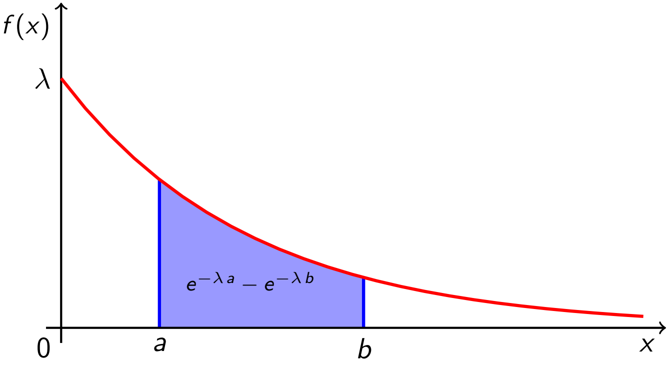
Exponential Distribution: Getting a Feeling

Interactive Visualisation (website link) or Python Notebook.

Exponential Distribution: Computing Probabilities

As with all continuous random variables, probabilities are found by integrating the density function over the interval of interest.

$$X \sim Exp(\lambda) \quad \Rightarrow \quad \mathbb{P}(a \leq X \leq b) = \int_a^b \lambda\, e^{-\lambda x}\, dx \;=\; e^{-\lambda\, a} - e^{-\lambda\, b}.$$



CDF and Other Properties

- The **cumulative distribution function** for $X \sim Exp(\lambda)$ is:

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0.$$

This follows immediately from the previous slide, by considering the interval $[a, b] = [0, x]$ (since $F(x) = \mathbb{P}(0 \leq X \leq x)$).

- Thus, we also get: $\mathbb{P}(X \geq x) = e^{-\lambda x}$. (from $1 - (1 - e^{-\lambda x})$)

- **Mean and Variance**:

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \mathbb{V}\text{ar}\,[X] = \frac{1}{\lambda^2}.$$

Note: The standard deviation of an $X \sim Exp(\lambda)$ is equal to its mean.

Example: Exponential – Toll Bridge

Suppose that the length of time (in minutes) between cars approaching a particular barrier at a toll bridge is exponentially distributed with mean 2. What is the probability that at least 3 minutes pass without a car approaching the barrier?

Solution

T: time between two consecutive cars.

$$T \sim Exp(\lambda) \;\Rightarrow\; \mathbb{E}[T] = \frac{1}{\lambda}. \qquad \text{So, } \mathbb{E}[T] = 2 \;\Rightarrow\; \lambda = \frac{1}{2}.$$

Since $T \sim Exp(1/2)$, we have

$$\mathbb{P}(T \geq t) = e^{-\lambda t} = e^{-t/2}.$$

Hence:

$$\mathbb{P}(T \geq 3) = e^{-3/2} = 0.223\,.$$

An Important Link with the Previous Chapter

- In the toll bridge example: on average, 1 car arrives every 2 minutes.
- This is a rate of $\lambda = 0.5$ cars per minute.
- Sounds familiar?

  $\rightarrow$ This is exactly a Poisson process with rate $\lambda = 0.5$.
- So the two views are linked:
  - Waiting time between cars: **Exponential(λ)**.
  - Number of cars in a time interval of length t: **Poisson(λt)**.

Connection: Poisson and Exponential

We can show the link. Suppose events follow a Poisson process with rate λ. Let $T =$ time until the next event.

- The event $T > t$ means: **no cars in the interval** $[0, t]$.
- But the number of cars in the interval $[0, t]$ is Poisson(λt). So:

$$\mathbb{P}(T > t) \;=\; \mathbb{P}(X_t = 0) \;=\; e^{-\lambda t}\,\frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

$$\implies \quad \mathbb{P}(T \le t) = 1 - e^{-\lambda t}$$

Summary: If events happen as a Poisson process with rate λ, then the waiting time between two events is Exponential(λ).

Example: Server log-ons

The number of user logins to a server can be modelled as a Poisson process with a mean of 25 logins per hour. What is the probability that there are no logins in a 6 minute interval? What is the probability that the time until the next login is between 2 and 3 minutes?

Solution

- $\lambda = \frac{25}{60} = \frac{5}{12}$ logins per minute.
- Call T: times between two consecutive logins. Then $T \sim Exp(\lambda)$.
- Prob of no logins in 6 minutes: $\mathbb{P}(T > 6) = e^{-\lambda 6} = e^{-5/2} = 0.082$.
- $$\begin{aligned} \mathbb{P}(2 \leq T \leq 3) &= e^{-\lambda 2} - e^{-\lambda 3} \\ &= e^{-5/6} - e^{-5/4} = 0.148. \end{aligned}$$

The Normal Distribution

The Normal Distribution: Motivation

The Normal distribution is one of the most widely used distributions. It naturally shows up in a variety of (very different) contexts:

- Medicine, biology, financial modelling, social and demographic sciences, ecology, climate, technology, ...

We will see one mathematical reason behind its ubiquity in so many contexts.

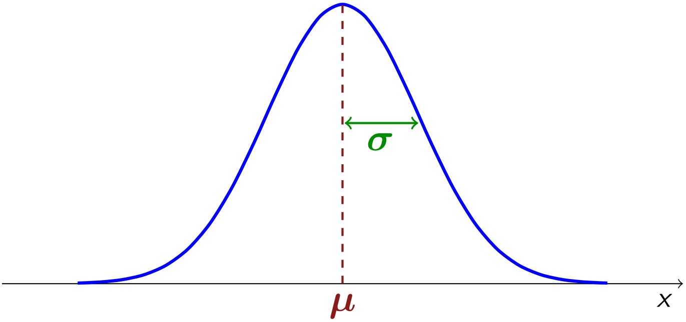
Needless to say, and precisely because of its countless applications, it plays a crucial role in:

- Statistics: Estimation, Hypothesis testing, Regression, ...
- Machine Learning: several algorithms (PCA, Clustering, Classification, ...)

The Normal Distribution

The Normal (or Gaussian) distribution is often referred to as the "bell-shaped" distribution. That is because of the shape of its density, which depends on two parameters:

- μ: the mean of the distribution
- σ^2: the variance of the distribution.



The Normal Distribution: Interactive Visualisation

Let's move to this link (or this link) to visualise interactively how the parameters μ and σ affect the shape of the distribution.

- The parameter μ (mean) affects the location of the density.
- The parameter σ (standard deviation) affects the spread of the density.
- The distribution is *symmetric*: mean, median and mode are all equal to μ.
- The parameter σ also affects the height of the density, since the area under the curve must be 1.

The Normal Density

If X is a **Normal random variable with mean μ and variance σ^2**, we write:

$$X \sim N(\mu, \sigma^2).$$

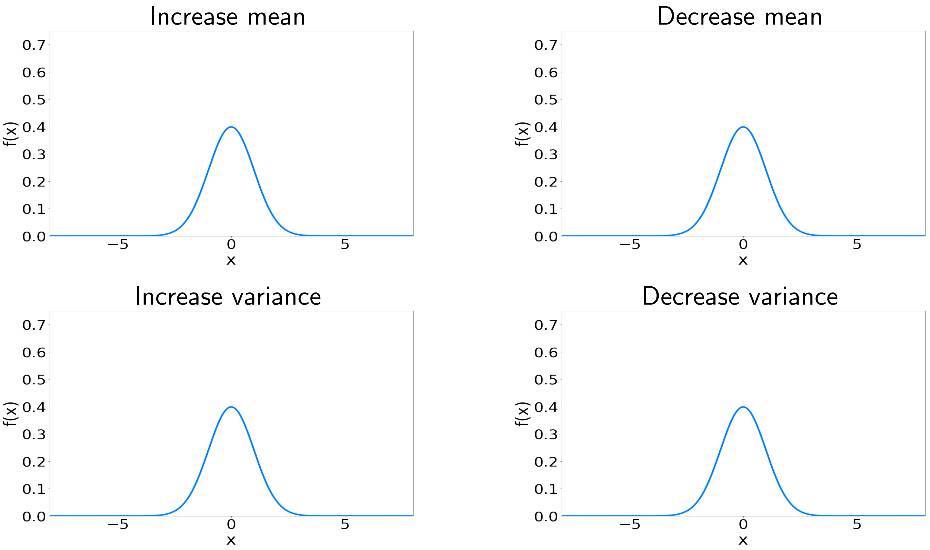
Note: The second argument is the value of the *variance*, not the standard deviation:

- X Normal with mean 3 and standard deviation 2 $\longrightarrow$ $X \sim N(3, 4)$
- Y Normal with mean -1 and standard deviation 5 $\longrightarrow$ $Y \sim N(-1, 25)$

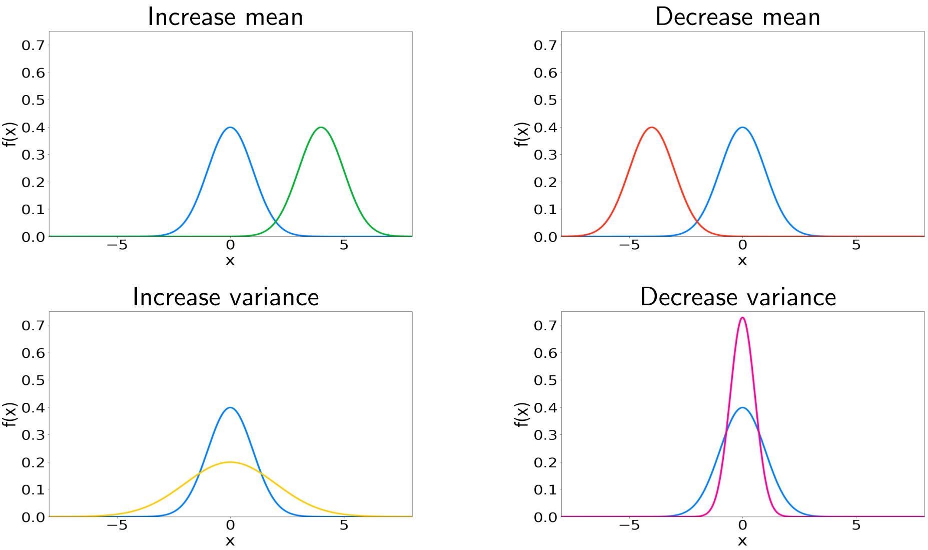
The functional form of the **Normal density** is as follows.

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

The Normal Distribution: Static Visualisation



The Normal Distribution: Static Visualisation



The Standard Normal Distribution

A Normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$ is called a **standard Normal distribution**. It is often denoted by Z:

$$Z \sim N(0, 1).$$

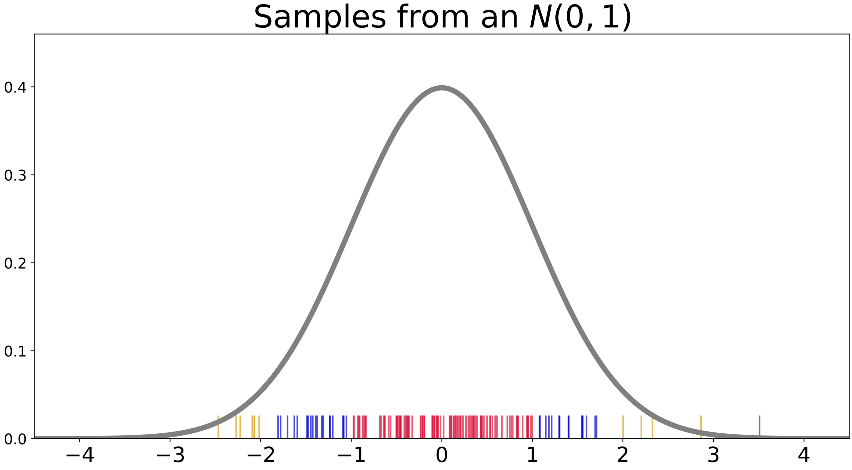
Its **density** is:

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2} \qquad \text{(set } \mu = 0 \text{ and } \sigma = 1 \text{ in equation (1))}$$

As we will see, any other Normal random variable can be expressed in terms of a standard Normal.

Python Notebook

Let's move to a Python Notebook to explore typical samples of an $N(0,1)$.



Standard Normal: Probability Computations

To compute $\mathbb{P}(a \leq Z \leq b)$, we would need to integrate f between a and b:

$$\mathbb{P}(a \leq Z \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-z^2/2}\, dz$$

However, **there is no closed-form expression** for the above definite integral:

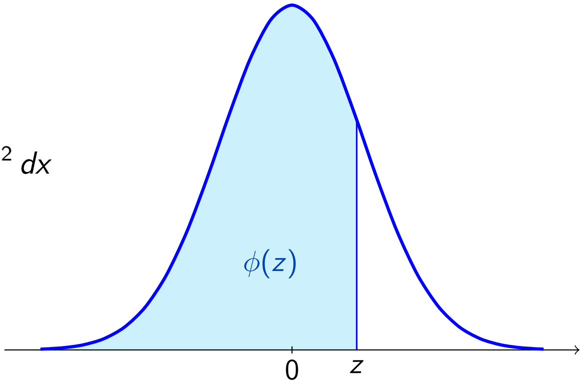
- Given two general values for a and b, there is no analytical formula to exactly compute the area under the Normal curve, between a and b.

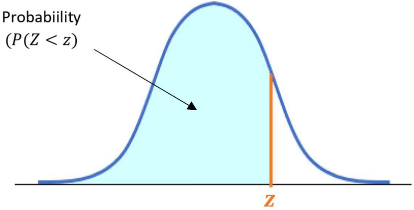
Instead, approximations are obtained using numerical methods. These are tabulated.

Standard Normal: Probability Computations

The standard Normal tables report the value of the cumulative distribution function (commonly denoted $\phi(z)$ for the standard Normal) at different values of z.

$$\phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$$



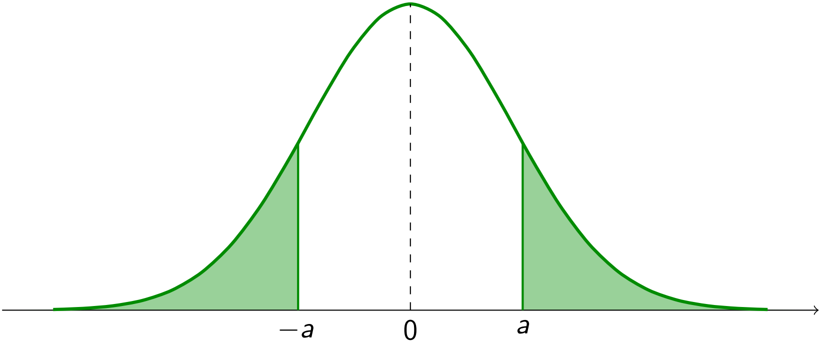


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5754
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.7	0.7580	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8079	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545

Standard Normal: Probability Computations

Note that the value of $\phi(z)$ is only reported for $z \geq 0$. This is because the standard Normal is symmetric around 0, hence:

$$\mathbb{P}(Z < -a) = \mathbb{P}(Z > a) \quad \forall\, a \geq 0.$$



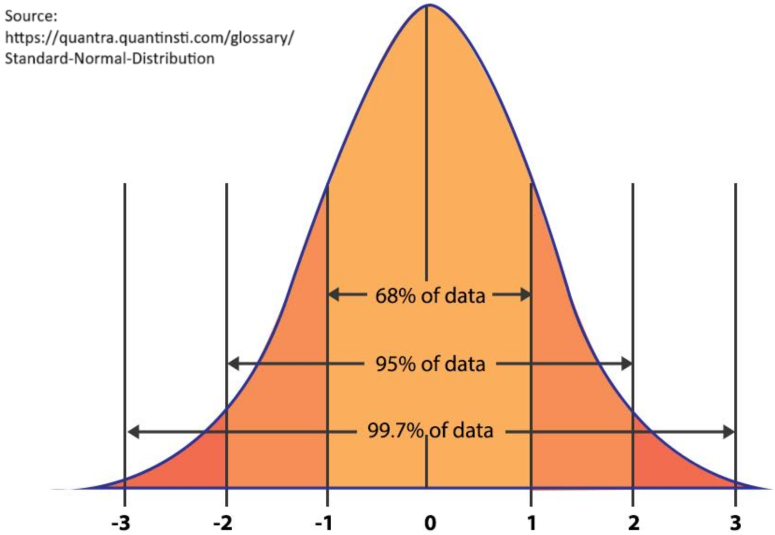
Examples: Standard Normal Computations

- $\mathbb{P}(Z < 1) = 0.8413$, directly from the table.
- $\mathbb{P}(Z < 0.56) = 0.7123$, directly from the table.
- $\mathbb{P}(1 \leq Z \leq 1.5) = \mathbb{P}(Z \leq 1.5) - \mathbb{P}(Z < 1) = 0.9332 - 0.8413 = 0.0919$.
- $\mathbb{P}(Z > 0.5) = 1 - \mathbb{P}(Z \leq 0.5) = 1 - 0.6915 = 0.3085$.
- $\mathbb{P}(Z < -1)$? That's the same as $\mathbb{P}(Z > 1)$. So:

$$\mathbb{P}(Z < -1) = \mathbb{P}(Z > 1) = 1 - \mathbb{P}(Z \leq 1) = 1 - 0.8413 = 0.1587.$$

- $\mathbb{P}(-1 \leq Z \leq 1) = \mathbb{P}(Z \leq 1) - \mathbb{P}(Z < -1) = 0.8413 - 0.1587 = 0.6826$.

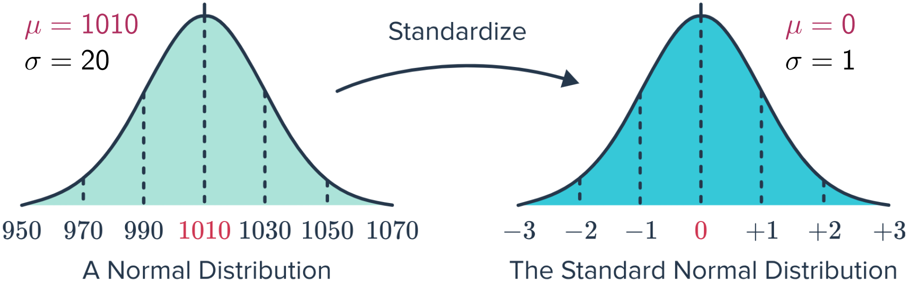
Standard Normal Distribution



What about a general Normal RV?

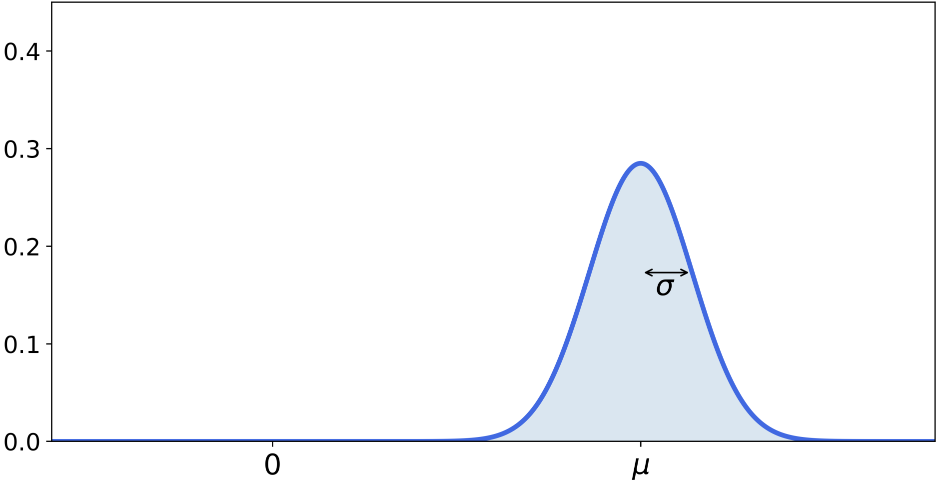
- What if we want to compute $\mathbb{P}(a \leq X \leq b)$, for a general Normal RV X? (*e.g.*, $X \sim N(1010, 20^2)$)

- We can easily "standardise" X, so that X becomes a standard Normal.

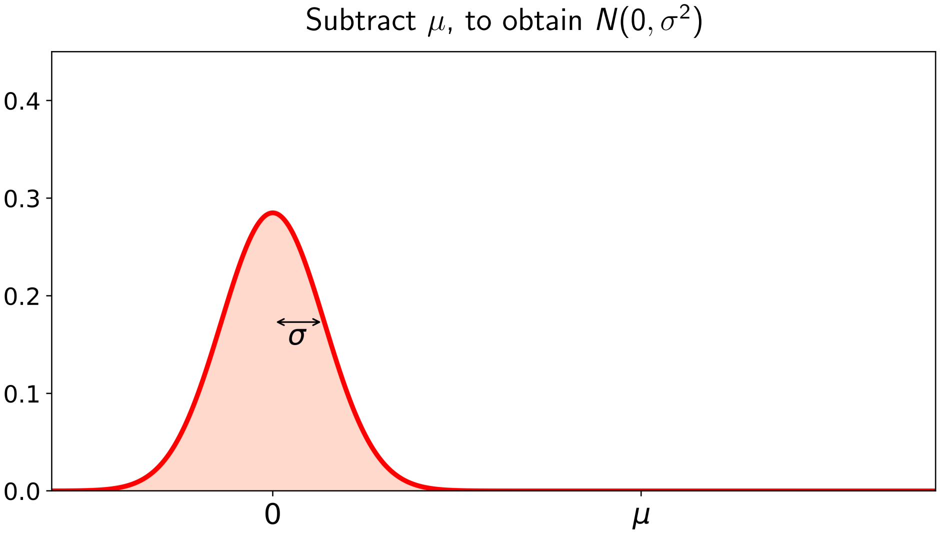


What about a general Normal RV? Visualisation

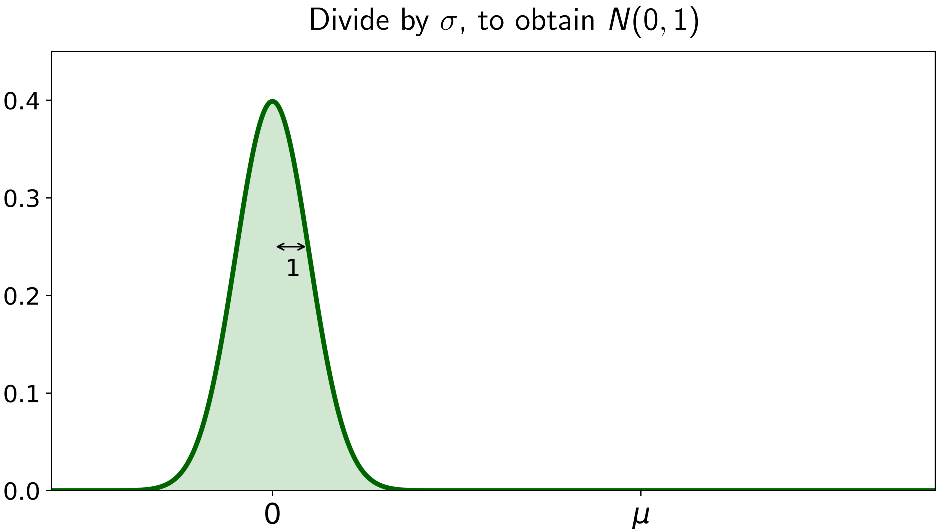
Start with $N(\mu, \sigma^2)$



What about a general Normal RV? Visualisation



What about a general Normal RV? Visualisation



Standardisation

Note 1: The above transformation is made of two steps:

1. First X is *centred* ($X \to X - \mu$), so that the new mean is 0.
2. Then, $X - \mu$ is divided by σ, so that the new standard deviation is 1.

$$X \sim N(\mu, \sigma^2) \quad \Rightarrow \quad Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

Note 2: The process can be also done in reverse order:

$$Z \sim N(0, 1) \;\; \Rightarrow \;\; X = \sigma Z + \mu \sim N(\mu, \sigma^2).$$

Standardisation

- The process we have just described (start from a random variable, subtract its mean, divide by its standard deviation) is called **standardisation**.
- It can be done on any RV, and will always return a rv with mean 0 and standard deviation 1.
- However, if starting from a *Normal* random variable, it will return another *Normal* random variable.

Standardise Single Values

- We can also standardise single observations coming from a normal distribution $N(\mu, \sigma^2)$. These are called ***z*-scores**.

Normal $N(\mu, \sigma^2)$		**Standard Normal** $N(0, 1)$
x	$\longrightarrow$	$z = \dfrac{x - \mu}{\sigma}$
$\mu + z\sigma$	$\longleftarrow$	z

- If the z-score for x is 2, then we know that x is 2 standard deviations to the right of the mean.

- Very useful to understand how "(a)typical" a value x is (for a $N(\mu, \sigma^2)$).

Probability Computations with Any Normal Distribution

- The previous two formulas allow us to compute probabilities for any Normal RV X:

$$\mathbb{P}(a \leq X \leq b), \quad \text{when } X \sim N(\mu, \sigma^2)$$

- How? Simply "rescale" the interval $[a, b]$ for X into the appropriate interval for $Z \sim N(0, 1)$.

- See next slide for the maths. But first, this link for an easier intuition! (remember to tick the "Show standard Normal curve" box, then play with x_1 and x_2)

Computation examples

1. Let $X \sim N(50, 10^2)$. What's $\mathbb{P}(40 \leq X \leq 70)$?

   We can write $X = 50 + 10\,Z$. Then:

   $$\begin{aligned}\mathbb{P}(40 \leq X \leq 70) &= \mathbb{P}(40 \leq 50 + 10\,Z \leq 70)\\ &= \mathbb{P}(-10 \leq 10\,Z \leq 20)\\ &= \mathbb{P}(-1 \leq Z \leq 2) = 0.8186.\end{aligned}$$

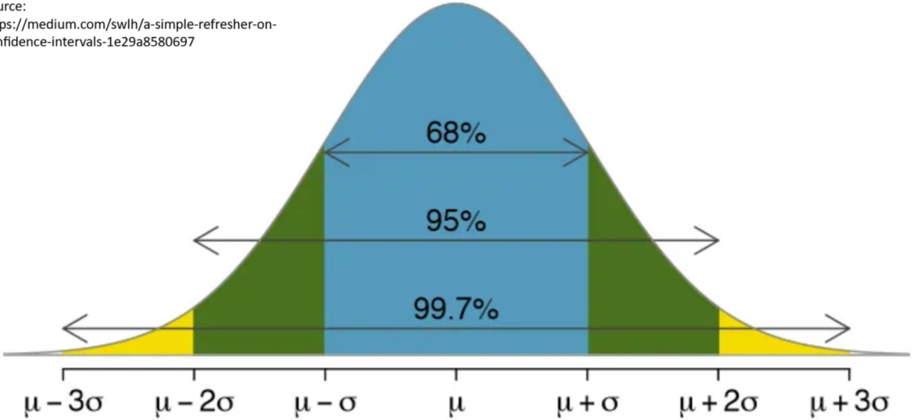
   Since X is centred in 50, and has a std of 10, falling between 40 and 70 means falling between -1 and $+2$ standard deviations from the mean.

2. Similarly, if $X \sim N(5, 0.4^2)$, then $(X - 5)/0.4 \sim N(0, 1)$. Hence:

   $$\mathbb{P}(X \leq 5.6) = \mathbb{P}\left(\frac{X-5}{0.4} \leq \frac{5.6-5}{0.4}\right) = \mathbb{P}(Z \leq 1.5) = 0.9332.$$

Standard Normal Distribution

Source:
https://medium.com/swlh/a-simple-refresher-on-confidence-intervals-1e29a8580697



Further Examples

Suppose $X \sim N(\mu = 5, \sigma^2 = 4)$. Then: $X = 5 + 2Z$, with $Z \sim N(0,1)$.

- Calculate $\mathbb{P}(X \leq 7)$.

$$\mathbb{P}(X \leq 7) = \mathbb{P}\left(5 + 2Z \leq 7\right) = \mathbb{P}(Z \leq 1) = 0.8413.$$

- Calculate $\mathbb{P}(X \geq 6)$.

$$\mathbb{P}(X \geq 6) = \mathbb{P}\left(5 + 2Z \geq 6\right) = \mathbb{P}(Z \geq 0.5) = 0.3085.$$

- Calculate $\mathbb{P}(X \leq 2)$.

$$\mathbb{P}(X \leq 2) = \mathbb{P}(5 + 2Z \leq 2) = \mathbb{P}(Z \leq -1.5) = 0.0668.$$

- Calculate $\mathbb{P}(2.5 \leq X \leq 7)$

$$\begin{aligned}\mathbb{P}(2.5 \leq X \leq 7) &= \mathbb{P}(2.5 \leq 5 + 2Z \leq 7) \\ &= \mathbb{P}(-2.5 \leq 2Z \leq 2)\end{aligned}$$

Example: Paper Friction

The coefficient of friction between sheets in a photocopier paper stack is studied. This is assumed to be Normally distributed with mean 0.55 and standard deviation 0.013. During system operation, the coefficient is measured at randomly selected times.

1. Find the probability that the coefficient falls between 0.53 and 0.56.

Let $Z = \frac{X-0.55}{0.013}$. Then $Z \sim N(0,1)$. Hence:

$$\begin{aligned}
\mathbb{P}(0.53 \leq X \leq 0.56) &= \mathbb{P}\left(\frac{0.53-0.55}{0.013} \leq \frac{X-0.55}{0.013} \leq \frac{0.56-0.55}{0.013}\right) \\
&= \mathbb{P}(-1.538 \leq Z \leq 0.769) \\
&= \mathbb{P}(Z \leq 0.769) - \mathbb{P}(Z \leq -1.538) \\
&= \mathbb{P}(Z \leq 0.769) - \mathbb{P}(Z > 1.538) \\
&= 0.7794 - 0.0618 = 0.7176.
\end{aligned}$$

Example: Paper Friction

The coefficient of friction between sheets in a photocopier paper stack is studied. This is assumed to be Normally distributed with mean 0.55 and standard deviation 0.013. During system operation, the coefficient is measured at randomly selected times.

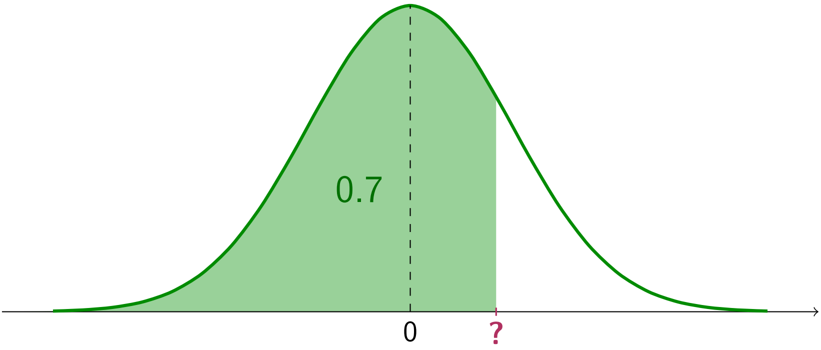
1. Find the probability that the coefficient falls between 0.53 and 0.56.
2. Is it likely to observe a friction coefficient below 0.52? Explain.

$$\begin{aligned}\mathbb{P}(X \leq 0.52) &= \mathbb{P}(Z \leq -2.308) \\ &= \mathbb{P}(Z > 2.308) \\ &= 1 - \mathbb{P}(Z < -2.308) = 0.0105.\end{aligned}$$

Hence, it is quite unlikely ($\approx 1\%$) to observe a friction coefficient below 0.52.

The Normal Distribution: Quantiles

- It can also be useful to use the Normal tables to find quantiles.
- Let's concentrate on the **standard Normal** first. What is the quantile of order 0.7? And the one of order 0.1?



The Normal Distribution: Quantiles

- We can use standard Normal tables to find quantiles for other Normals.
- Suppose $X \sim N(2, 9)$. For what value a is $\mathbb{P}(X \leq a) = 0.1$?

First, from the tables for the standard Normal, we find that the quantile of order 0.1 for an $N(0, 1)$ is $a = -1.28$:

$$\mathbb{P}(Z \leq -1.28) = 0.1\,.$$

Then, we have:

$$\text{For an } N(0,1): \quad a = -1.2816$$

$$\text{For an } N(2,3^2): \quad 2 + 3a = -1.8448$$

Example: Exam Results

Scores on an examination are assumed to be Normally distributed with a mean of 78 and a variance of 36.

1. Suppose that students scoring in the top 10% of this distribution are to receive "Excellent" as grade. What is the minimum score a student must achieve to obtain an A grade?
2. What must be the cutoff for passing the exam if the examiner wants only the lowest 25% of all scores to fail?
3. Find, approximately, what proportion of students have scores 5 or more points above the passing cutoff.

Solution: Exam Results

1. We need to find the 90th percentile of the grade distribution, *i.e*:

$$a \in \mathbb{R} \text{ such that } \mathbb{P}(X \leq a) = 0.9, \text{ when } X \sim N(78, 6^2).$$

From the standard Normal tables, we find $a = 1.2816$.

Since $X = 78 + 6Z$, the 90th percentile for X is:

$$78 + 6 \times 1.2816 = 85.69\,.$$

Solution: Exam Results

2. We need to find the grade a for which $\mathbb{P}(X \leq a) = 0.25$.

   Again, from the standard Normal tables, we find

$$\mathbb{P}(Z \leq -0.675) = 0.25.$$

   Through the transformation $X = 78 + 6Z$, we transform the standardised value -0.675 to the following grade for X:

$$78 + 6 \times (-0.675) = 73.95\,.$$

Solution: Exam Results

3. We are asked to find the probability:

$$\mathbb{P}(X \geq 73.95 + 5).$$

Writing X as $X = 78 + 6\,Z$, we get:

$$\begin{aligned}\mathbb{P}(X \geq 78.95) &= \mathbb{P}(78 + 6\,Z \geq 78.95)\\ &= \mathbb{P}(Z \geq 0.16) = 1 - \mathbb{P}(Z \leq 0.16)\\ &= 1 - 0.5636 = 0.4364.\end{aligned}$$